

NGDA Dataset Report

Official NGDA Title: 114th Congressional District

Metadata Record Title: TIGER/Line Shapefile, 2014, Series Information For The 114th Congressional District National Shapefile

A-16 NGDA Theme: Governmental Units, and Administrative and Statistical Boundaries

Executive NGDA Theme Champion(s):

Name: Timothy F. Trainor

Agency: DOC-Census

Email: timothy.f.trainor@census.gov

Theme Lead(s):

Name: Lynda Liptrap

Agency: Census

Email: lynda.a.liptrap@census.gov

Dataset Manager(s):

Name: Dierdre Bevington Attardi

Agency: Census

Email: Dierdre.Bevington.Attardi@census.gov

Metadata:

Registration Status: Complete

Registered on 1/16/2015

GeoPlatform Link*: <http://www.geoplatform.gov/node/243/f081c116-0570-455d-9178-943d6b3dbf91>

Data.gov Metadata Link*: <http://catalog.data.gov/harvest/object/8025685a-293d-4b71-a969-b9741627f22e/html>

*If the metadata has been updated and reharvested after publication of this report, the link may no longer be valid. The dataset may be searched for manually in Data.gov or GeoPlatform.gov.

NGDA Lifecycle Maturity Assessment (LMA) Report

Time Frame:

Baseline assessment responses include dataset activities from 1998 to 2015.

LMA Submission:

Status: Complete

Date: 9/30/2015

Extension Requested: Yes

LMA Reviewer(s):

Supervisor: James Whitehorne, Chief of the Census Redistricting and Voting Rights Data Office

Theme Lead: Lynda A Liptrap, Chief, Federal Geographic Coordination Branch

Executive Champion: Did not review

SAOGI*: Did not review

Other: Did not review

LMA Verifier:

Name: Lynda A Liptrap, James Whitehorne

Email: Lynda.A.Liptrap@census.gov, James.Whitehorne@census.gov

Attachments:

To get access to any attachments referenced in the report, email the LMA Help Desk at NGDA_LMA_help@fgdc.gov. Please use the subject "Dataset Report Attachment(s)" and indicate the associated official NGDA title.

*Senior Agency Official for Geospatial Information (SAOGI)

Lifecycle Maturity Assessment (LMA) Summary

Overall Maturity:

Optimized; Established

General Questions: 100%

Optimized; Established

Stage 1 - Define/Plan: 100%

Optimized; Established

Stage 2 - Inventory/Evaluate: 100%

Optimized; Established

Stage 3 - Obtain: 100%

Optimized; Established

Stage 4 - Access: 100%

Optimized; Established

Stage 5 - Maintain: 100%

Optimized; Established

Stage 6 - Use/Evaluate: 100%

Optimized; Established

Stage 7 - Archive: 100%

Optimized; Established

NGDA Dataset Maturity Definitions:

How To Calculate Maturity: https://www.geoplatform.gov/sites/default/files/How_to_Calculate_Maturity.pdf

Maturity	Maturity Characteristics for All Lifecycle Stages
Optimized; Established Rank = 5	Dataset meets virtually all business needs of all users. The dataset is considered authoritative by owners and secondary users. It is curated across all stages of the approved lifecycle. Future needs are defined on a regular basis and resources for addressing both current and future business requirements are available.
Mature; Consistent Rank = 4	Dataset meets all the business needs of the primary owner and most of the secondary users. The dataset is curated and used as authoritative by the primary owner. Dataset is used widely by secondary users actively engaged in sustaining the dataset. Future needs are identified and steps are planned to address these. All stages are supported and reviewed on a recurring basis. The dataset is well managed in relation to the approved lifecycle.
Managed; Predictable Rank = 3	Dataset meets a significant number of the business needs of the primary owner and is widely used as an authoritative resource by secondary users. Benchmark activities are occurring in at least four of the approved lifecycle stages. Management practices in relation to the approved lifecycle is moderate but consistent. Dataset is integrating changing business requirements in lifecycle stages impacting overall maturity.
Transition; Transformation Rank = 2	Dataset meets business needs of the primary owner and has moderate use by secondary users. Benchmark activities are occurring in at least three stages. Efforts to integrate funding, include partners, and obtain data are not supported in a sustained manner. Management practices in relation to the stages of the approved lifecycle is limited.
Planned; Initial Development Rank = 1	Dataset limited in meeting business needs of the primary owner. Benchmark activities in the approved lifecycle are just starting to consider secondary uses, partnerships are forming to support additional dataset uses. Dataset development is in a very early stage. Minimal or limited management against the benchmarks in the approved lifecycle.
No Activity Rank = no activity	Dataset meets project or local business needs of the primary owner, secondary or additional uses or users were not considered, not recognized as an authoritative data or is part of a similar dataset. Not managed to any of the benchmarks in the approved lifecycle.

General Questions for All Stages

1) Is there a recurring process to obtain funding for all lifecycle stages of this dataset?

Answer: Funding support is part of agency budget on a recurring basis, funding is consistent and tied to business processes, and supports all lifecycle stages.

Justification Comment:

Attachment(s): 0

This funding comes from the annual Redistricting project for the Redistricting Data Office (RDO), Geography Division (GEO), and Population Division (POP).

This is annual funding and is appropriated each year for running these programs.

The funding fluctuates based on the Redistricting Program (RDP) business cycle but is included in the annual budgets for the work necessary.

No interagency funding needed as this is conducted solely through the Census Bureau with the State's participation.

2) Is there a process in place to ensure that open government and transparency guidelines are followed in all lifecycle stages for this dataset?

Answer: Process is published as appropriate with respect to sensitivity requirements, process is transparent, published appropriately.

Justification Comment:

Attachment(s): 0

Collection of these boundaries is done in an open process with all states having more than one congressional district being contacted for their input. The materials for participation are provided to the participants and posted to the public website for the office.

The metadata is published to public portals for discovery.

There are also a number of external and internal standards that apply to this data, such as:

Section 508 of the Workforce Investment Act of 1998 mandates that when developing, procuring, maintaining, or using electronic and information technology, each Federal department or agency, including the United States Postal Service, shall ensure, unless an undue burden would be imposed on the department or agency, that the electronic and information technology allows, regardless of the type of medium of the technology.

-individuals with disabilities who are Federal employees to have access to and use of information and data that is comparable to the access to and use of the information and data by Federal employees who are not individuals with disabilities; and

-individuals with disabilities who are members of the public seeking information or services from a Federal department or agency to have access to and use of information and data that is comparable to the access to and use of the information and data by such members of the public who are not individuals with disabilities

The following links provide information on the data security issues and confidentiality protocol:

-Bureau of the Census Handbook for Information Technology Security located at

<http://cww.census.gov/po/ims/revpolhandbk2.pdf>

-Title 13 U.S.C. OMB No. 0607-0809 located at <http://cww.census.gov/po/legislate/13usc.htm>

Numbered Memorandum G9 Guidelines for "Census Confidential" Materials located at

<http://cww.census.gov/po/ims/nmgen9.htm>

3) Are there processes and tools in place so that staff are sufficiently knowledgeable to ensure a continuity of the dataset for all stages of the lifecycle, especially during staffing transitions?

Answer: Processes and tools to ensure dataset continuity are in place and implemented for all lifecycle stages.

Justification Comment:

Attachment(s): 0

Staff roles and responsibilities are defined and there is ongoing, informal training such as seminars on Census geography, data, and products. In general, the maintenance of the boundary and feature data

contained in the TIGER/Line Shapefiles are delegated to the subject matter experts (SMEs) for each geographic boundary or feature. These SMEs follow their own respective data maintenance procedures, which include identifying key personnel and assigning tasks for data maintenance, projects and training.

The Software Requirements Specifications (SRS) document for the TIGER/Line Shapefiles is maintained and updated with each iteration of the public release of the shapefiles.

The TIGER/Line technical documentation exists and is maintained. It can be viewed at:

http://www2.census.gov/geo/pdfs/maps-data/data/tiger/tgrshp2014/TGRSHP2014_TechDoc.pdf

STAGE 1 - Define/Plan

4) Are user and business requirements defined and formalized?

Answer: A recurring process is in place, including defining new partner and stakeholder business needs as they arise, and is fully implemented.

Justification Comment:

Attachment(s): 0

The Census Redistricting Data Office conducts a thorough review of the Redistricting Data Program (RDP) with the States, the primary stakeholders, at the conclusion of each decennial.

During this review, the business requirements for collection of the redistricting data program's geographies is discussed, defined, modified, and formalized in a publication titled "The View from the States":

<http://www.census.gov/rdo/pdf/DesignPL94-171.pdf>

The RDP is conducted in response to legislation, [P.L. 94-171], that directs the Census Bureau to provide states with the small area geography necessary for conducting state legislative redistricting. Over time this has proven to be congressional, state legislative, and voting districts along with the ability to influence each census' tabulation blocks.

5) How are partners/stakeholders involved in the requirements collection process?

Answer: A recurring process is in place, including defining new partner and stakeholder business needs as they arise, and is fully implemented.

Justification Comment:

Attachment(s): 0

Input from the state partners helps define what is to be collected and on what frequency. They are consulted after each decennial for feedback on the dataset and their feedback is incorporated in the review document "The View from the States".

The Geography Division subject matter experts comprise the internal partners, stakeholders and stewards of the geospatial products. All subject matter experts, including those who work directly with the public users, regularly contribute to the dataset and resource requirements and maintenance of the dataset. Roles and responsibilities, including which office is responsible for submitting requirements for the data, are clearly defined through the program-specific workflow charts and the Geography Division Management Information System (GEO MIS).

6) Is there a quality assurance process for the dataset?

Answer: Quality assurance published as appropriate with respect sensitivity requirements.

Justification Comment:

Attachment(s): 0

All public products are reviewed with manual and automated quality assessment procedures prior to release. In addition, source data from partnerships must pass accuracy tests before being integrated with TIGER.

All QA processes are documented including instructions to the QA analysts, notes and progress reports. The QA requirements are reviewed during each iteration of production and are updated as needed.

In addition, a copy of shapefiles and block equivalency files are provided to the partners for review and acceptance. Once they accept the plan they are provided a form to complete to certify their acceptance.

7) Is there a process to evaluate the sensitivity, privacy, and confidentiality of this dataset?

Answer: Sensitivity, privacy, and confidentiality evaluations fully implemented, reviewed and updated on a recurring basis.

Justification Comment:

Attachment(s): 0

All of the geospatial data products are Title 13 compliant, including shapefiles and relationship files that contained address ranges. The address range subject matter experts employ data suppression techniques in those data records with unique and/or single-range addresses. All files are reviewed through quality assurance checks prior to every public release to ensure that the data are Title 13 compliant.

8) Are defined data standards used in collecting, processing, and/or rendering the data?

Answer: Standards fully implemented documented and published as appropriate.

Justification Comment:

Attachment(s): 0

There are defined standards in that the collection of these districts uses census geography and identifies any faces that are split and where that split occurs.

Internal database imbedded business rules are used to standardize processes. Standardized procedures are used for rendering cartographic products and web services.

The data format is image files, and shapefiles.

The geographic concepts, record layouts, metadata attributes, values and definitions for rendering TIGER/Line shapefiles are synthesized in the TIGER/Line Shapefile technical documentation and shapefile metadata. They are also documented in "The View from the States".

STAGE 2 - Inventory/Evaluate

9) Is there a process for determining if data necessary to meet requirements already exist from other sources (either within or outside the agency) before collecting or acquiring new data?

Answer: Process for determining appropriate data is being reused fully implemented, reviewed, and updated on a regular basis.

Justification Comment:

Attachment(s): 0

There is no formal research necessary to identify internal or external datasets for the creation of this dataset. These datasets are provided by the states using our defined procedures such as submitting whole block and split block plans, population listings, and split block lists and images.

STAGE 3 - Obtain

10) Is there a process for obtaining data in relation to this dataset?

Answer: Process is fully implemented, reviewed and updated on a regular basis.

Justification Comment:

Attachment(s): 0

The collection of Congressional Districts is for a data that is owned by the states and we are reflecting those boundaries within our MAF/TIGER database. This is a partnership driven data collection using the collection plan established by the Redistricting Data Office.

This data is collected on a voluntary basis and is provided freely. Since this is a reflection of the state's data using the state's interpretation of how their boundaries align with Census geography, no transformation is needed.

This process to obtain the data involves direct consultation with the partners, the states, to reproduce their data in TIGER. There are defined procedures for the hand-off of these plans.

11) Is the metadata in a FGDC endorsed geospatial metadata standard?

Answer: Metadata is available in a format endorsed by the FGDC, it fully describes the dataset and provides all the information required to make the dataset discoverable, accessible, and usable.

Justification Comment:

Attachment(s): 0

Metadata is published and available to the public for the TIGER/Line shapefile product. The metadata is in the ISO 19115 format and is discoverable through federal data portals and packaged with the downloadable files. The metadata template is specified in internal Census requirements documents for the public TIGER/Line Shapefiles.

12) How complete is the geographic coverage as defined in the requirements for the dataset?

Part 1 Answer: Business requirements for cyclic updates identified and a process is in place.

Part 2 Answer: Dataset has presently attained the greatest geographic coverage as defined by the current requirements or roughly 100%.

Justification Comment:

Attachment(s): 0

The TIGER/Line Shapefiles cover the entire United States, the District of Columbia, Puerto Rico, and the Island areas, which is the defined coverage area.

STAGE 4 - Access

13) Do you have a process for providing users access to the data in an open digital machine readable format?

Answer: User access process is fully implemented, data is available, process is reviewed and updated on a recurring basis.

Justification Comment:

Attachment(s): 0

The Redistricting Data Office conducts a multiphase Redistricting Data Program.

One phase of that program is the official data delivery as required by P.L. 94-171 to the states.

This phase of the program identifies the individuals in each state who are official recipients, and provides that data as the official delivery.

This process is documented in the decennial publication known as "The View from the States" and in the Federal Register notice announcing the Redistricting data program.

In preparing "The View" stakeholder concerns and comments are solicited and worked into the process. In addition, all of the data from this program, congressional, state legislative, and voting districts are provided through the Bureau's public digital dissemination process.

There is an established process by which data is made available on the Census website for download in a machine-readable format. This includes the shapefiles, metadata, and technical documentation.

Partners and the public may also view the data through maps and online geography viewers.

Access may be obtained at: <http://www.census.gov/geography.html>

STAGE 5 - Maintain

14) Is there a maintenance process for updating and storing the dataset?

Answer: Dataset maintenance process is fully implemented and processes are reviewed and periodically updated.

Justification Comment:

Attachment(s): 0

The Census Redistricting Data Office conducts periodic calls for updates to the boundaries of products generated for the redistricting data program.

TIGER, to include the congressional district data, is stored in the MAF-TIGER Database (MTDB).

Business rules, geographic area edits and core Application Program Interface (API) rules are in place to maintain the geography. The business rules and core API reduce the chance that geography is altered in error. The business rules are well-documented and updated when new geography or geographic relationships are added. The versioning of the data is well organized by the benchmarking

process and creation of benchmark specific shapefiles. There is documentation for the business rules, geographic area edits, core API, benchmarking, and database/products vintages, which is Census Bureau versioning. The data are typically updated continuously through various programs and operations.

15) Is there an error correction process as part of dataset maintenance?

Answer: Error correction process includes user notification, process reviewed on a recurring basis.

Justification Comment:

Attachment(s): 0

Any feedback from external users which result in change to the TIGER/Line Shapefiles are documented through the production processes of the next shapefile release. If a large-scale error or errors are found after undergoing quality assurance, the files are removed from the website, a note is posted about the error, the error is logged into a tracking database, and new files may be created and released. The timing of the correction and release of new data varies based on the situation.

STAGE 6 - Use/Evaluate

16) Is there a process to determine if the dataset meets user needs?

Answer: Process is fully implemented and repeated on a recurring basis.

Justification Comment:

Attachment(s): 0

The Census Redistricting Data Office conducts an official evaluation of the program which includes the products generated for the program and their usefulness. This feedback is integrated into the program design to improve both the data quality and usability to the primary stakeholders. The results of this evaluation and the recommendations are included in the publication "The View from the States" published between each census.

17) Is there a process to provide users information on how to access and properly use the dataset?

Answer: Process is fully implemented supporting access and proper use, process is reviewed on a recurring basis.

Justification Comment:

Attachment(s): 0

We do have a release plan annually for the TIGER/Line shapefiles and this dataset. The plan has been in use since 2007 and it covers how the files are named, zipped, organized on the ftp site, and how the web interface works. The plan covers how the files will look and will be organized once they are released.

The TIGER/Line Shapefiles Technical documentation and metadata provides documentation for the access to and the use of the data.

The TIGER Geodatabase Record Layouts are available at:

http://www2.census.gov/geo/pdfs/maps-data/data/tiger/tgrshp2014/2014_TIGER_GDB_Record_Layouts.pdf

The TIGER/Line Shapefile Technical Documentation is available at:

http://www2.census.gov/geo/pdfs/maps-data/data/tiger/tgrshp2014/TGRSHP2014_TechDoc.pdf

There are no access constraints to the public users of the data; however, the Geography Division posts the following caveat in the metadata files, technical documentation and on the product website: The TIGER/Line Shapefile products are not copyrighted however, TIGER/Line and Census TIGER are registered trademarks of the U.S. Census Bureau. These products are free to use in a product or publication, however acknowledgement must be given to the U.S. Census Bureau as the source.

18) Are the business processes and management practices assessed to meet changing technology?

Answer: Assessment process is fully implemented for taking advantage of changing technology, process is reviewed on a recurring basis.

Justification Comment:

Attachment(s): 0

The business processes and management practices are evaluated each decade, using input from the

states and using the most current Census technologies deployed to continually update them. This involves work with the external partners and technological innovations incorporated into other Census operations and maintenance.

STAGE 7 - Archive

19) Is there an archiving process for the dataset?

Answer: Archival and disposition processes are fully implemented.

Justification Comment:

Attachment(s): 0

We do archive copies of the files including the old disk copies of the files. Creating a disk copy is part of our requirements for the product and part of the plan for removing the files from the internal GEO servers when they remove the files from the place on the server where they are created. Our archive copies contain all the data that are on the ftp site, but it is not stored in the same directory structure on our archived disks due to the way the software creates the files on our server.

Census maintains CD/DVD copies of all vintages of TIGER created to date. We create our own copies (currently on DVD). We have at least one copy of each disk.

In addition, 2010 TIGER/Line data were archived by NARA.